SOME NEW OBSERVATIONS ON INTERPOLATION IN THE SPECTRAL UNIT BALL

GAUTAM BHARALI

Abstract. We present several results associated to a holomorphic-interpolation problem for the spectral unit ball $\Omega_n,\ n \geq 2$. We begin by showing that a known necessary condition for the existence of a $\mathcal{O}(\mathbb{D};\Omega_n)$-interpolant ($\mathbb{D}$ here being the unit disc in $\mathbb{C}$), given that the matricial data are non-derogatory, is not sufficient. We provide next a new necessary condition for the solvability of the two-point interpolation problem – one which is *not* restricted only to non-derogatory data, and which incorporates the Jordan structure of the prescribed data. We then use some of the ideas used in deducing the latter result to prove a Schwarz-type lemma for holomorphic self-maps of $\Omega_n,\ n \geq 2$.

1. Introduction and Statement of Results

The interpolation problem referred to in the title, and which links the assorted results of this paper, is the following ($\mathbb{D}$ here will denote the open unit disc centered at $0 \in \mathbb{C}$):

(*) *Given M distinct points $\zeta_1, \dots, \zeta_M \in \mathbb{D}$ and matrices $W_1, \dots, W_M$ in the spectral unit ball $\Omega_n := \{W \in M_n(\mathbb{C}) : r(W) < 1\}$, find conditions on $\{\zeta_1, \dots, \zeta_M\}$ and $\{W_1, \dots, W_M\}$ such that there exists a holomorphic map $F : \mathbb{D} \longrightarrow \Omega_n$ satisfying $F(\zeta_j) = W_j,\ j = 1, \dots, M$.*

In the above statement, $r(W)$ denotes the spectral radius of the $n \times n$ matrix W. Under a very slight simplification – i.e. that the interpolant F in (*) is required to satisfy $\sup_{\zeta\in\mathbb{D}} r(F(\zeta)) < 1$ – the paper [5] provides a characterisation of the interpolation data $((\zeta_1, W_1), \dots, (\zeta_M, W_M))$ that admit an interpolant of the type described. However, this characterisation involves a non-trivial search over a region in $\mathbb{C}^{n^2 M}$. Thus, there is interest in finding alternative characterisations that either: *a)* circumvent the need to perform a search; or *b)* reduce the dimension of the search-region. In this regard, a new idea idea was introduced by Agler & Young in the paper [1]. This idea was further developed over several works – notably in [2], in the papers [7] and [8] by Costara, and in David Ogle's thesis [13]. It can be summarised in two steps as follows:

- If the matrices $W_1, \dots, W_M$ are all non-derogatory, then (*) is equivalent to an interpolation problem in the *symmetrized polydisc* $G_n,\ n \geq 2$, which is defined as

$$G_n := \left\{(s_1, \dots, s_n) \in \mathbb{C}^n : \text{all the roots of } z^n + \sum\nolimits_{j=1}^{n} (-1)^j s_j z^{n-j} = 0 \text{ lie in } \mathbb{D}\right\}.$$

1991 *Mathematics Subject Classification.* Primary: 30E05, 47A56; Secondary: 32F45.

Key words and phrases. Complex geometry, Carathéodory metric, minimial polynomial, Schwarz lemma, spectral radius, spectral unit ball.

This work is supported in part by a grant from the UGC under DSA-SAP, Phase IV.

To appear in ***Integral Eqns. Operator Theory***.

- The G_n-interpolation problem is shown to share certain aspects of the classical Nevanlinna-Pick problems, either by establishing conditions for a von Neumann inequality for $\overline{G}_n$ – note that $\overline{G}_n$ is compact – or through function theory.

It would be useful, at this stage, to recall the following

Definition 1.1. A matrix $A \in M_n(\mathbb{C})$ is said to be *non-derogatory* if the geometric multiplicity of each eigenvalue of A is 1 (regardless of its algebraic multiplicity). The matrix A being non-derogatory is equivalent to A being similar to the companion matrix of its characteristic polynomial – i.e., if $z^n + \sum_{j=1}^n s_j z^{n-j}$ is the characteristic polynomial then

$$A \text{ is non-derogatory} \iff A \text{ is similar to } \begin{bmatrix} 0 & & & -s_n \\ 1 & 0 & & -s_{n-1} \\ & \ddots & \ddots & \vdots \\ \mathbf{0} & & 1 & -s_1 \end{bmatrix}_{n\times n}.$$

The Agler-Young papers treat the case $n = 2$, while the last two works cited above consider the higher-dimensional problem. The reader is referred to [2] for a proof of the equivalence of (*), given non-derogatory matricial data, and the appropriate G_n-interpolation problem. The similarity condition given in Definition 1.1 is central to establishing this equivalence.

Before presenting the first result of this paper, we need to examine what is known about (*) from the perspective of the G_n-interpolation problem. Since we would like to focus on the matricial interpolation problem, we will paraphrase the results from [13] and [8] in the language of non-derogatory matrices. Given an $n \times n$ complex matrix W, let its characteristic polynomial $\chi^W(z) = z^n + \sum_{j=1}^n (-1)^j s_j(W) z^{n-j}$, and define the rational function

$$\mathsf{f}(z;W) := \frac{\sum_{j=1}^n j s_j(W)(-1)^j z^{j-1}}{\sum_{j=0}^{n-1}(n-j)s_j(W)(-1)^j z^j}.$$

Then, the most general statement that is known about (*) is:

Result 1.2 (paraphrased from [13] and [8])**.** *Let $\zeta_1, \dots, \zeta_M$ be M distinct points in $\mathbb{D}$ and let $W_1, \dots, W_M \in \Omega_n$ be non-derogatory matrices. If there exists a map $F \in \mathcal{O}(\mathbb{D}, \Omega_n)$ such that $F(\zeta_j) = W_j$, $j = 1, \dots, M$, then the matrices*

$$\left[\frac{1 - \overline{\mathsf{f}(z;W_j)}\mathsf{f}(z;W_k)}{1 - \overline{\zeta_j}\zeta_k}\right]_{j,k=1}^M \geq 0 \quad \text{for each } z \in \overline{\mathbb{D}}. \tag{1.1}$$

Here, and elsewhere in this paper, given two complex domains X and Y, $\mathcal{O}(X;Y)$ will denote the class of all holomorphic maps from X into Y.

Remark 1.3. The matrices in (1.1) may appear different from those in [13, Corollary 5.2.2], but the latter are, in fact, $*$-congruent to the matrices above.

Even though Result 1.2 provides only a necessary condition, (1.1) is more tractable for small values of M than the Bercovici-Foias-Tannenbaum condition. Its viability as a sufficient condition, at least for small M, has been discussed in both [13] and [8]. This is reasonable because *the latter condition is sufficient* when $n = 2$ and $M = 2$ (and the given matrices are, of course, non-derogatory); see [4]. Given all these developments, it seems appropriate to begin with the following:

Observation 1.4. *When $n \geq 3$, the condition* (1.1) *is* not sufficient *for the existence of a $\mathcal{O}(\mathbb{D};\Omega_n)$-interpolant for the prescribed data $((\zeta_1, W_1), \dots, (\zeta_M, W_M))$, where each $W_j \in \Omega_n,\ j = 1, \dots, M,$ is non-derogatory.*

The above observation relies on ideas from complex geometry; specifically – estimates for invariant metrics on the symmetrized polydisc $G_n,\ n \geq 3$. Our argument follows from a recent study [11] of the Carathéodory metric on $G_n,\ n \geq 3$. This argument is presented in the next section.

Observation 1.4 takes us back to the drawing board when it comes to realising goals of the type *(a)* or *(b)* (as in the opening paragraph) to determine whether a $\mathcal{O}(\mathbb{D};\Omega_n)$-interpolant exists for a given data-set. Thus, new conditions that are inequivalent to (1.1) are desirable for the same reasons as those offered in [2] and [3]. To wit: all extant approaches to implementing the Bercovici-Foias-Tannenbaum solution of (*) are computational, and rely upon various search algorithms. Rigorous analytical results, even if they only indicate when a data-set $((\zeta_1, W_1), \dots, (\zeta_M, W_M))$ does *not* admit an $\mathcal{O}(\mathbb{D};\Omega_n)$ interpolant – i.e. necessary conditions – provide tests of existing algorithms/software and illustrate the complexities of (*). We will say more about this; but first – notations for our next result. Given $z_1, z_2 \in \mathbb{D}$, the *pseudohyperbolic distance* between these points, written $\mathcal{M}_{\mathbb{D}}(z_1, z_2)$, is defined as:

$$\mathcal{M}_{\mathbb{D}}(z_1, z_2) \ := \ \left|\frac{z_1 - z_2}{1 - \overline{z_2}z_1}\right| \quad \forall z_1, z_2 \in \mathbb{D}$$

We can now state our next result.

Theorem 1.5. *Let $F \in \mathcal{O}(\mathbb{D};\Omega_n)$, $n \geq 2$, and let $\zeta_1, \zeta_2 \in \mathbb{D}$. Write $W_j = F(\zeta_j)$, and let*

$$\begin{aligned}\sigma(W_j) \ &:= \ \textit{the } \text{set} \textit{ of eigenvalues of } W_j,\ j = 1, 2\\ &\qquad \textit{(i.e. elements of } \sigma(W_j) \textit{ are } \text{not} \textit{ repeated according to multiplicity).}\end{aligned}$$

If $\lambda \in \sigma(W_j)$, then let $m(\lambda)$ denote the multiplicity of λ as a zero of the minimal polynomial of W_j. Then:

(1.2)

$$\max\left\{\max_{\mu\in\sigma(W_2)}\prod_{\lambda\in\sigma(W_1)}\mathcal{M}_{\mathbb{D}}(\mu,\lambda)^{m(\lambda)},\ \max_{\lambda\in\sigma(W_1)}\prod_{\mu\in\sigma(W_2)}\mathcal{M}_{\mathbb{D}}(\lambda,\mu)^{m(\mu)}\right\} \ \leq \ \left|\frac{\zeta_1 - \zeta_2}{1 - \overline{\zeta_2}\zeta_1}\right|.$$

Referring back to our previous paragraph: one could ask whether Theorem 1.5 is able to highlight any complexities of (*) that Result 1.2 misses. There are two parts to the answer:

1) *The Jordan structure of the data-set $((\zeta_1, W_1), (\zeta_2, W_2))$:* Several well-known examples from [6] and [2] reveal that the existence of a $\mathcal{O}(\mathbb{D};\Omega_n)$-interpolant, $n \geq 2$, is sensitive to the Jordan structure of the matrices $W_1, \dots, W_M$. However, to the best of our knowledge, there are no results in the literature to date that incorporate information on the Jordan structures or the minimal polynomials of $W_1, \dots, W_M$. In contrast, the following example shows that information on minimal polynomials is vital – i.e. that with the correct information about the minimal polynomials of $F(\zeta_1)$ and $F(\zeta_2)$, condition (1.2) is sharp.

Example 1.6. For $n \geq 3$ and $d = 2, \dots, n-1$, define the holomorphic map $F_d : \mathbb{D} \longrightarrow \Omega_n$ by

$$F_d(\zeta) := \left[\begin{array}{cccc|c} 0 & & & \zeta & \\ 1 & 0 & & 0 & \\ & \ddots & \ddots & \vdots & 0 \\ & & 1 & 0 & \\ \hline & & 0 & & \zeta\mathbb{I}_{n-d} \end{array}\right]_{n \times n}, \qquad \zeta \in \mathbb{D},$$

where $\mathbb{I}_{n-d}$ denotes the identity matrix of dimension $n-d$ for $1 < d < n$. Let $\zeta_1 = 0$ and $\zeta_2 = \zeta$. One easily computes – in the notation of Theorem 1.5 – that:

$$\begin{aligned} \max_{\mu \in \sigma(W_2)} \prod\nolimits_{\lambda \in \sigma(W_1)} \mathcal{M}_{\mathbb{D}}(\mu, \lambda)^{m(\lambda)} &= |\zeta|, \\ \max_{\lambda \in \sigma(W_1)} \prod\nolimits_{\mu \in \sigma(W_2)} \mathcal{M}_{\mathbb{D}}(\lambda, \mu)^{m(\mu)} &= |\zeta|^2, \end{aligned}$$

where the first equality holds because W_1 is nilpotent of order d. So, (1.2) is satisfied as an equality for the given choice of ζ_1 and ζ_2 – which is what was meant above by saying that (1.2) *is sharp*. $\square$

2) *Comparison with* (1.1)*:* Theorem 1.5 would not be effective in testing any of the existing algorithms used in the implementation of the Bercovici-Foias-Tannenbaum solution to (*) if (1.1) were a universally stronger necessary condition than (1.2). However, (1.1) is devised with non-derogatory data in mind, whereas no *simple* interpolation condition was hitherto known for pairs of arbitrary matrices in Ω_n. Hence, by choosing any one of W_1 and W_2 to be derogatory, one would like to examine how (1.1) and (1.2) compare. This leads to our next observation.

Observation 1.7. *For each $n \geq 3$, we can find a data-set $((\zeta_1, W_1), (\zeta_2, W_2))$ for which* (1.2) *implies that it cannot admit any $\mathcal{O}(\mathbb{D}; \Omega_n)$-interpolant, whereas* (1.1) *provides no information.*

An example pertinent to this observation is presented at the end of Section 3. As for Theorem 1.5, it may be viewed as a Schwarz lemma for mappings between $\mathbb{D}$ and the spectral unit ball. Note that the inequality (1.2) is preserved under automorphisms of $\mathbb{D}$ and under the "obvious" automorphisms of Ω_n (the full automorphism group $\mathsf{Aut}(\Omega_n)$, $n \geq 2$, is not known). The proof of Theorem 1.5 is presented in Section 3.

The key new idea in the proof of Theorem 1.5 – i.e. to focus on the minimal polynomial of certain crucial matrices that lie in the range of F – pays off in obtaining a result that is somewhat removed from the our main theme. The result in question is a generalisation of the following theorem of Ransford and White [14, Theorem 2]:

$$(1.3) \qquad G \in \mathcal{O}(\Omega_n; \Omega_n) \text{ and } G(0) = 0 \implies r(G(X)) \leq r(X) \;\; \forall X \in \Omega_n.$$

One would like to generalise (1.3) in the way the Schwarz-Pick lemma generalises the Schwarz lemma for $\mathbb{D}$ – i.e. by formulating an inequality that is valid without assuming that the holomorphic mapping in question has a fixed point. This generalisation is as follows:

Theorem 1.8. *Let $G \in \mathcal{O}(\Omega_n;\Omega_n)$, $n \geq 2$, and define $d_G :=$ the degree of the minimal polynomial of $G(0)$. Then:*

$$r(G(X)) \leq \frac{r(X)^{1/d_G} + r(G(0))}{1 + r(G(0))r(X)^{1/d_G}} \quad \forall X \in \Omega_n. \tag{1.4}$$

Furthermore, the inequality (1.4) *is sharp in the sense that there exists a non-empty set $\mathfrak{S}_n \subset \Omega_n$ such that given any $A \in \mathfrak{S}_n$ and $d = 1, \dots, n$, we can find a $\mathsf{G}^{A,d} \in \mathcal{O}(\Omega_n;\Omega_n)$ such that*

$$d_{\mathsf{G}^{A,d}} = d, \quad \textit{and}$$

$$r(\mathsf{G}^{A,d}(A)) = \frac{r(A)^{1/d} + r(\mathsf{G}^{A,d}(0))}{1 + r(\mathsf{G}^{A,d}(0))r(A)^{1/d}}. \tag{1.5}$$

2. A Discussion of Observation 1.4

We begin this discussion with a couple of definitions from complex geometry. Given a domain $\Omega \subset \mathbb{C}^n$, the *Carathéodory pseudodistance* between two points $z_1, z_2 \in \Omega$ is defined as

$$c_\Omega(z_1, z_2) := \sup\{p_{\mathbb{D}}(f(z_1), f(z_2)) : f \in \mathcal{O}(\Omega;\mathbb{D})\},$$

where $p_{\mathbb{D}}$ is the Poincaré distance on $\mathbb{D}$ (and $p_{\mathbb{D}}$ is given by $p_{\mathbb{D}}(\zeta_1,\zeta_2) = \tanh^{-1}(\mathcal{M}_{\mathbb{D}}(\zeta_1,\zeta_2))$ for $\zeta_1, \zeta_2 \in \mathbb{D}$). In the same setting, the *Lempert functional* on $\Omega \times \Omega$, is defined as

$$\begin{aligned} &\widetilde{\kappa}_\Omega(z_1, z_2) \\ &:= \inf\{p_{\mathbb{D}}(\zeta_1,\zeta_2) : \exists \psi \in \mathcal{O}(\mathbb{D};\Omega) \text{ and } \zeta_1,\zeta_2 \in \mathbb{D} \text{ such that } \psi(\zeta_j) = z_j,\ j = 1, 2.\}. \end{aligned} \tag{2.1}$$

It is not hard to show that the set on the right-hand side above is non-empty. The reader is referred to Chapter III of [10] for details. Next, we examine a few technical objects. For the remainder of this section, $S = (s_1, \dots, s_n)$ will denote a point in $\mathbb{C}^n$, $n \geq 2$. For $z \in \overline{\mathbb{D}}$ define the rational map $f_n(z;S) := (\widetilde{s}_1(z;S), \dots, \widetilde{s}_{n-1}(z;S))$, $n \geq 2$, by

$$\widetilde{s}_j(z;S) := \frac{(n-j)s_j - z(j+1)s_{j+1}}{n - zs_1}, \quad S \in \mathbb{C}^n \text{ s.t. } n - zs_1 \neq 0, \quad j = 1, \dots, (n-1).$$

Next, define

$$F(Z;\bullet) := f_2(z_1;\bullet) \circ \cdots \circ f_n(z_{n-1};\bullet) \quad \forall Z = (z_1, \dots, z_{n-1}) \in \overline{\mathbb{D}}^{n-1},$$

where the second argument varies through that region in $\mathbb{C}^n$ where the right-hand side above is defined. The connection of these objects with our earlier discussions is established via

$$f(z;S) := \frac{\sum_{j=1}^{n} js_j(-1)^j z^{j-1}}{\sum_{j=0}^{n-1}(n-j)s_j(-1)^j z^j}, \quad z \in \overline{\mathbb{D}},$$

and S varies through that region in $\mathbb{C}^n$ where the right-hand side above is defined. Note the resemblance of $f(z;S)$ to $\mathsf{f}(z;W)$ defined earlier. From Theorem 3.5 of [8], we excerpt:

Result 2.1. *Let $S = (s_1, \dots, s_n)$ denote a point in $\mathbb{C}^n$. Then:*

1) $f(z;\bullet) = F(z, \dots, z;\bullet)\ \forall z \in \overline{\mathbb{D}}$, *wherever defined.*
2) $S \in G_n$ *if and only if* $\sup_{z \in \overline{\mathbb{D}}} |f(z;S)| < 1$, $n \geq 2$.

3) *If $S \in G_n,\ n \geq 2$, then*

$$\sup_{z\in\overline{\mathbb{D}}} |f(z;S)| \;=\; \sup_{Z\in\overline{\mathbb{D}}^{n-1}} |F(Z;S)|.$$

For convenience, let us refer to the Carathéodory pseudodistance on $G_n,\ n \geq 2$, by c_n. Next, define – here we refer to Section 2 of [11] – the following distance function on G_n

$$\mathfrak{p}_n(S,T) \;:=\; \max_{Z\in(\partial\mathbb{D})^{n-1}} p_{\mathbb{D}}(F(Z;S), F(Z;T)) \quad \forall S,T \in G_n. \tag{2.2}$$

This is the distance function – whose properties have been studied in [11] – we shall exploit to support Observation 1.4. The well-definedness of the right-hand side above follows from parts (2) and (3) of Result 2.1 above. Furthermore, since $F(Z;S), F(Z;T) \in \mathbb{D}$ for each $Z \in \overline{\mathbb{D}}^{n-1}$ whenever $S,T \in G_n,\ n \geq 2$, it follows simply from the definition that

$$c_n(S,T) \;\geq\; \mathfrak{p}_n(S,T) \quad \forall S,T \in G_n. \tag{2.3}$$

Since we have now adopted certain notations from [11], we must make the following

Note. *We have opted to rely on the notation of* [8]*. This leads to a slight discrepancy between our definition of $\mathfrak{p}_n$ in* (2.2) *and that in* [11]*. This discrepancy is easily reconciled by the observation that $F(\bullet;S)$ used here and in* [8] *will have to be read as $F(\bullet;-s_1,s_2,\dots,(-1)^n s_n)$ in* [11]*. This is harmless because $S \in G_n \iff (-s_1,s_2,\dots,(-1)^n s_n) \in G_n$.*

Let us now refer back to the condition (1.1) with $M = 2$. An easy calculation involving 2×2 matrices reveals that

$$\textit{When } M=2,\ (1.1) \iff \sup_{z\in\overline{\mathbb{D}}}\left|\frac{\mathsf{f}(z;W_1)-\mathsf{f}(z;W_2)}{1-\overline{\mathsf{f}(z};W_2)\mathsf{f}(z;W_1)}\right| \leq \left|\frac{\zeta_1-\zeta_2}{1-\overline{\zeta_2}\zeta_1}\right|.$$

If W_1 is nilpotent of order n (recall that all matrices occuring in (1.1) are non-derogatory), then $\mathsf{f}(\bullet;W_1)\equiv 0$. Of course, $W_2 \in \Omega_n$ implies that $(s_1(W_2),\dots,s_n(W_2)) \in G_n$. By part (2) of Result 2.1, $\mathsf{f}(z;W_2) \in \mathbb{D}\ \forall z \in \overline{\mathbb{D}}$. This leads to the following key fact:

(2.4) *When $M = 2$ and W_1 is nilpotent of order n,*

$$(1.1) \iff \sup_{z\in\overline{\mathbb{D}}} \tanh^{-1}|\mathsf{f}(z;W_2)| = \sup_{z\in\overline{\mathbb{D}}} p_{\mathbb{D}}(0,\mathsf{f}(z;W_2)) \leq p_{\mathbb{D}}(\zeta_1,\zeta_2).$$

We now appeal to Proposition 2 in [11], i.e. $\mathfrak{p}_n(0,\bullet) \neq c_n(0,\bullet)$ for each $n \geq 3$. Let us now fix $n \geq 3$. Let $S_0 \in G_n\setminus\{0\}$ be such that $c_n(0,S_0) > \mathfrak{p}_n(0,S_0)$. Let $\varepsilon_0 > 0$ be such that $c_n(0,S_0) = \mathfrak{p}_n(0,S_0) + 2\varepsilon_0$. Let us write $S_0 = (s_{0,1},\dots,s_{0,n})$ and choose two matrices $W_1, W_2 \in \Omega_n$ as follows:

$$W_1 = \textit{a nilpotent of order } n, \quad W_2 = \begin{bmatrix} 0 & & & (-1)^{n-1}s_{0,n} \\ 1 & 0 & & (-1)^{n-2}s_{0,n-1} \\ & \ddots & \ddots & \vdots \\ 0 & & 1 & s_{0,1} \end{bmatrix}_{n\times n},$$

i.e. W_2 is the companion matrix of the polynomial $z^n + \sum_{j=1}^n (-1)^j s_{0,j} z^{n-j}$. We emphasize the following facts that follow from this choice of W_1 and W_2

$$\mathsf{f}(\cdot, W_1) = f(\cdot; 0, \dots, 0) \equiv 0, \ \ \mathsf{f}(\cdot, W_2) = f(\cdot; S_0), \tag{2.5}$$

W_1 *and* W_2 *are, by construction, non-derogatory.*

The relations in (2.5) are cases of a general correspondence between matrices in Ω_n and points in G_n , given by the surjective, holomorphic map $\Pi_n : \Omega_n \longrightarrow G_n$, where

$$\Pi_n(W) \ := \ (s_1(W), \dots, s_n(W)),$$

and $s_j(W), \ j = 1, \dots n$, are as defined in the beginning of this article.

Let us pick two distinct points $\zeta_1, \zeta_2 \in \mathbb{D}$ such that

$$p_{\mathbb{D}}(\zeta_1, \zeta_2) - \varepsilon_0 \ < \ \mathfrak{p}_n(0, S_0) \ \leq \ p_{\mathbb{D}}(\zeta_1, \zeta_2). \tag{2.6}$$

Assume, now, that (1.1) is a *sufficient condition* for the existence of a $\mathcal{O}(\mathbb{D}; \Omega_n)$-interpolant. Then, in view of the choices of W_1, W_2, the second inequality in (2.6), and (2.5) we get

$$\sup_{z \in \overline{\mathbb{D}}} \tanh^{-1} |\mathsf{f}(z; W_2)| \ = \ \sup_{z \in \partial\mathbb{D}} \tanh^{-1} |\mathsf{f}(z; W_2)| \ = \ \mathfrak{p}_n(0, S_0) \ \leq \ p_{\mathbb{D}}(\zeta_1, \zeta_2). \tag{2.7}$$

The first equality in (2.7) is a consequence of part (2) of Result 2.1: since $S_0 \in G_n$, the rational function $\mathsf{f}(\cdot; W_2) = f(\cdot; S_0) \in \mathcal{O}(\mathbb{D}) \bigcup \mathcal{C}(\overline{\mathbb{D}})$, whence the equality follows from the Maximum Modulus Theorem. But now, owing to the equivalence (2.4), the estimate (2.7) implies, by assumption, that there exists an interpolant $F \in \mathcal{O}(\mathbb{D}; \Omega_n)$ such that $F(\zeta_j) = W_j, \ j = 1, 2$. Then, $\Pi_n \circ F : \mathbb{D} \longrightarrow G_n$ satisfies $\Pi_n \circ F(\zeta_1) = 0$ and $\Pi_n \circ F(\zeta_2) = S_0$. Then, by the definition of the Lempert functional (for convenience, we denote the Lempert functional of G_n by $\widetilde{\kappa}_n$)

$$\begin{aligned} \widetilde{\kappa}_n(0, S_0) \ \leq \ p_{\mathbb{D}}(\zeta_1, \zeta_2) \ &< \ \mathfrak{p}_n(0, S_0) + \varepsilon_0 && \text{(from (2.6), 1st part)} \\ &< \ c_n(0, S_0). && \text{(by definition of } \varepsilon_0\text{)} \end{aligned}$$

But, for any domain Ω, the Carathéodory pseudodistance and the Lempert function always satisfy $c_\Omega \leq \widetilde{\kappa}_\Omega$. Hence, we have just obtained a contradiction. Hence our assumption that (1.1) is sufficient for the existence of an $\mathcal{O}(\mathbb{D}, \Omega_n)$-interpolation, for $n \geq 3$, must be false.

3. The Proof of Theorem 1.5

The proofs in this section depend crucially on a theorem by Vesentini. The result is as follows:

Result 3.1 (Vesentini, [15])**.** *Let $\mathcal{A}$ be a complex, unital Banach algebra and let $r(x)$ denote the spectral radius of any element $x \in \mathcal{A}$. Let $f \in \mathcal{O}(\mathbb{D}; \mathcal{A})$. Then, the function $\zeta \longmapsto r(f(\zeta))$ is subharmonic on $\mathbb{D}$.*

The following result is the key lemma of this section. The proof of Theorem 1.5 is reduced to a simple application of this lemma. The structure of this proof is reminiscent of [12, Theorem 1.1]. This stems from the manner in which Vesentini's theorem is used. The essence of the trick below goes back to Globevnik [9]. The reader will notice that Theorem 1.5 specialises to Globevnik's Schwarz lemma when $W_1 = 0$.

Lemma 3.2. *Let $F \in \mathcal{O}(\mathbb{D}; \Omega_n)$. For each $\lambda \in \sigma(F(0))$, define $m(\lambda)$:=the multiplicity of λ as a zero of the minimal polynomial of $F(0)$. Define the Blaschke product*

$$B(\zeta) \ := \prod_{\lambda \in \sigma(F(0))} \left(\frac{\zeta - \lambda}{1 - \overline{\lambda}\zeta} \right)^{m(\lambda)}, \quad \zeta \in D.$$

Then $|B(\mu)| \leq |\zeta| \ \forall \mu \in \sigma(F(\zeta))$.

Proof. The Blaschke product B induces a matrix function $\widetilde{B}$ on Ω_n: for any matrix $A \in \Omega_n$, we set

$$\widetilde{B}(A) \ := \prod_{\lambda \in \sigma(F(0))} (\mathbb{I} - \overline{\lambda}A)^{-m(\lambda)}(A - \lambda\mathbb{I})^{m(\lambda)},$$

which is well-defined on Ω_n because whenever $\lambda \neq 0$,

$$(\mathbb{I} - \overline{\lambda}A) \ = \ \overline{\lambda}(\mathbb{I}/\overline{\lambda} - A) \in GL(n, \mathbb{C}).$$

Furthermore, since $\zeta \longmapsto (\zeta - \lambda)/(1 - \overline{\lambda}\zeta)$, $|\lambda| < 1$, has a power-series expansion that converges uniformly on compact subsets of $\mathbb{D}$, it follows from standard arguments that

$$\sigma(\widetilde{B}(A)) \ = \ \{B(\mu) : \mu \in \sigma(A)\} \quad \text{for any } A \in \Omega_n. \tag{3.1}$$

By the definition of the minimal polynomial, $\widetilde{B} \circ F(0) = 0$. Since $\widetilde{B} \circ F(0) = 0$, there exists a holomorphic map $\Phi \in \mathcal{O}(\mathbb{D}; M_n(\mathbb{C}))$ such that $\widetilde{B} \circ F(\zeta) = \zeta\Phi(\zeta)$. Note that

$$\sigma(\widetilde{B} \circ F(\zeta)) \ = \ \sigma(\zeta\Phi(\zeta)) \ = \ \zeta\sigma(\Phi(\zeta)) \quad \forall \zeta \in \mathbb{D}. \tag{3.2}$$

Since $\sigma(\widetilde{B} \circ F(\zeta)) \subset \mathbb{D}$, the above equations give us:

$$r(\Phi(\zeta)) \ < \ 1/R \quad \forall \zeta : |\zeta| = R, \ R \in (0, 1). \tag{3.3}$$

Taking $\mathcal{A} = M_n(\mathbb{C})$ in Vesentini's theorem, we see that $\zeta \longmapsto r(\Phi(\zeta))$ is subharmonic on the unit disc. Applying the Maximum Principle to (3.3) and taking limits as $R \longrightarrow 1^-$, we get

$$r(\Phi(\zeta)) \ \leq \ 1 \quad \forall \zeta \in \mathbb{D}. \tag{3.4}$$

In view of (3.1), (3.2) and (3.4), we get

$$|B(\mu)| \ \leq \ |\zeta| r(\Phi(\zeta)) \ \leq \ |\zeta| \quad \forall \mu \in \sigma(F(\zeta)).$$

$\square$

We are now in a position to provide

3.3. The proof of Theorem 1.5. Define the disc automorphisms

$$M_j(\zeta) \ := \ \frac{\zeta - \zeta_j}{1 - \overline{\zeta_j}\zeta}, \quad j = 1, 2,$$

and write $\Phi_j = F \circ M_j^{-1}$, $j = 1, 2$. Note that $\Phi_1(0) = W_1$. For $\lambda \in \sigma(W_1)$, let $m(\lambda)$ be as stated in the theorem. Define the Blaschke product

$$B_1(\zeta) \ := \prod_{\lambda \in \sigma(W_1)} \left(\frac{\zeta - \lambda}{1 - \overline{\lambda}\zeta} \right)^{m(\lambda)}, \quad \zeta \in \mathbb{D}.$$

Applying Lemma 3.2, we get

$$\left|\frac{\zeta_1-\zeta_2}{1-\overline{\zeta_2}\zeta_1}\right| = |M_1(\zeta_2)| \geq \prod_{\lambda\in\sigma(W_1)}\left|\frac{\mu-\lambda}{1-\overline{\lambda}\mu}\right|^{m(\lambda)}$$
$$= \prod_{\lambda\in\sigma(W_1)} \mathcal{M}_{\mathbb{D}}(\mu,\lambda)^{m(\lambda)} \quad \forall\mu\in\sigma(\Phi_1(M_1(\zeta_2)))=\sigma(W_2). \tag{3.5}$$

Now, swapping the roles of ζ_1 and ζ_2 and applying the same argument to

$$B_2(\zeta) := \prod_{\mu\in\sigma(W_2)}\left(\frac{\zeta-\mu}{1-\overline{\mu}\zeta}\right)^{m(\mu)}, \quad \zeta\in\mathbb{D},$$

we get

$$\left|\frac{\zeta_1-\zeta_2}{1-\overline{\zeta_2}\zeta_1}\right| \geq \prod_{\mu\in\sigma(W_2)} \mathcal{M}_{\mathbb{D}}(\lambda,\mu)^{m(\mu)} \quad \forall\lambda\in\sigma(W_1). \tag{3.6}$$

Combining (3.5) and (3.6), we get

$$\max\left\{\max_{\mu\in\sigma(W_2)}\prod_{\lambda\in\sigma(W_1)}\mathcal{M}_{\mathbb{D}}(\mu,\lambda)^{m(\lambda)},\ \max_{\lambda\in\sigma(W_1)}\prod_{\mu\in\sigma(W_2)}\mathcal{M}_{\mathbb{D}}(\lambda,\mu)^{m(\mu)}\right\} \leq \left|\frac{\zeta_1-\zeta_2}{1-\overline{\zeta_2}\zeta_1}\right|.$$

□

We conclude this section with an example.

Example 3.4. *An illustration of Observation 1.7*

We begin by pointing out that the phenomenon below is expected for $n=2$. We want to consider $n>2$ and show that *there is no interpolant for the following data*, but that this cannot be inferred from (1.1). First the matricial data: let $n=2m,\ m\geq 2$, and let

$$W_1 = \textit{any block-diagonal matrix with two } m\times m\textit{-blocks that are each nilpotent of order } m. \tag{3.7}$$

Next, for an $\alpha\in\mathbb{D}$, $\alpha\neq 0$, let

$$W_2 = \textit{the companion matrix of the polynomial } (z^{2m}-\alpha z^m).$$

Note that, by construction, W_2 is non-derogatory. We have the characteristic polynomials $\chi^{W_1}(z)=z^m$ and $\chi^{W_2}(z)=z^{2m}-\alpha z^m$. Hence

$$\mathsf{f}(\cdot;W_1) \equiv 0, \qquad \mathsf{f}(z;W_2) = \frac{-m\alpha z^{m-1}}{2m-m\alpha z^m}.$$

We recall, from Section 2, the following equivalent form of (1.1):

$$\textit{When } M=2, \text{ (1.1)} \iff \sup_{z\in\overline{\mathbb{D}}}\left|\frac{\mathsf{f}(z;W_1)-\mathsf{f}(z;W_2)}{1-\overline{\mathsf{f}(z};W_2)\mathsf{f}(z;W_1)}\right| \leq \left|\frac{\zeta_1-\zeta_2}{1-\overline{\zeta_2}\zeta_1}\right|. \tag{3.8}$$

Since, clearly, $\mathsf{f}(\cdot;W_2)\in\mathcal{O}(\mathbb{D})\bigcap\mathcal{C}(\overline{\mathbb{D}})$, by the Maximum Modulus Theorem

$$\sup_{z\in\overline{\mathbb{D}}}\left|\frac{\mathsf{f}(z;W_1)-\mathsf{f}(z;W_2)}{1-\overline{\mathsf{f}(z};W_2)\mathsf{f}(z;W_1)}\right| = \sup_{z\in\partial\mathbb{D}}\frac{m|\alpha|}{|2m-m\alpha z^m|}$$
$$= \frac{m|\alpha|}{2m-m|\alpha|} < |\alpha|. \tag{3.9}$$

Observe that $\sigma(W_1) = \{0\}$ and $\sigma(W_2) = \{0, |\alpha|^{1/m}e^{i(2\pi j+\mathsf{Arg}(\alpha))/m},\ j = 1, \dots, m\}$. Therefore,

$$\max_{\mu\in\sigma(W_2)} \prod\nolimits_{\lambda\in\sigma(W_1)} \mathcal{M}_{\mathbb{D}}(\mu,\lambda)^{m(\lambda)} \ = \ |\alpha|,$$
$$\max_{\lambda\in\sigma(W_1)} \prod\nolimits_{\mu\in\sigma(W_2)} \mathcal{M}_{\mathbb{D}}(\lambda,\mu)^{m(\mu)} \ = \ 0.$$

We set $\zeta_1 = 0$ and pick $\zeta_2 \in \mathbb{D}$ in such a way that

$$\frac{m|\alpha|}{2m - m|\alpha|} \ < \ |\zeta_2| \ = \ \left|\frac{\zeta_1 - \zeta_2}{1 - \overline{\zeta_2}\zeta_1}\right| \ < \ |\alpha|. \tag{3.10}$$

Such a choice of ζ_2 is made possible by the inequality (3.9). In view of the last calculation above, we see that the data-set $((W_1, \zeta_1), (W_2, \zeta_2))$ constructed violates the inequality (1.2). Thus, there is no $\mathcal{O}(\mathbb{D}, \Omega_{2m})$-interpolant for this data-set. In contrast, since the equivalent form (3.8) of (1.1) is satisfied, the latter does not yield any information about the existence of a $\mathcal{O}(\mathbb{D}, \Omega_{2m})$-interpolant. ☐

4. The Proof of Theorem 1.8

In order to prove Theorem 1.8, we shall need the following elementary

Lemma 4.1. *Given a fractional-linear transformation $T(z) := (az + b)/(cz + d)$, if $T(\partial\mathbb{D}) \Subset \mathbb{C}$, then $T(\partial\mathbb{D})$ is a circle with*

$$\text{centre}(T(\partial\mathbb{D})) \ = \ \frac{b\overline{d} - a\overline{c}}{|d|^2 - |c|^2}, \qquad \text{radius}(T(\partial\mathbb{D})) \ = \ \frac{|ad - bc|}{\big||d|^2 - |c|^2\big|}.$$

We are now in a position to present

4.2. The proof of Theorem 1.8. Let $G \in \mathcal{O}(\Omega_n; \Omega_n)$ and let $\lambda_1, \dots, \lambda_s$ be the distinct eigenvalues of $G(0)$. Define $m(j)$:=the multiplicity of the factor $(\lambda - \lambda_j)$ in the minimal polynomial of $G(0)$. Define the Blaschke product

$$B_G(\zeta) \ := \ \prod_{j=1}^{s} \left(\frac{\zeta - \lambda_j}{1 - \overline{\lambda_j}\zeta}\right)^{m(j)}, \quad \zeta \in \mathbb{D}.$$

B_G induces the following matrix function which, by a mild abuse of notation, we shall also denote as B_G

$$B_G(Y) \ := \ \prod_{j=1}^{s} (\mathbb{I} - \overline{\lambda_j}Y)^{-m(j)}(Y - \lambda_j\mathbb{I})^{m(j)} \quad \forall Y \in \Omega_n,$$

which is well-defined on Ω_n precisely as explained in the proof of Lemma 3.2. Once again, owing to the analyticity of B_G on Ω_n,

$$\sigma(B_G(Y)) \ = \ \{B_G(\lambda) : \lambda \in \sigma(Y)\} \quad \forall Y \in \Omega_n,$$

whence $B_G : \Omega_n \longrightarrow \Omega_n$. Therefore, if we define

$$H(X) \ := \ B_G \circ G(X) \quad \forall X \in \Omega_n,$$

then $H \in \mathcal{O}(\Omega_n;\Omega_n)$ and, by construction, $H(0)=0$. By the Ransford-White result, $r(H(X)) \le r(X)$, or, more precisely

$$\max_{\mu\in\sigma(G(X))} \left\{ \prod_{j=1}^{s} \left| \frac{\mu-\lambda_j}{1-\overline{\lambda_j}\mu} \right|^{m(j)} \right\} \ \le\ r(X) \quad \forall X \in \Omega_n.$$

In particular:

$$\max_{\mu\in\sigma(G(X))} \left[\mathsf{dist}_{\mathcal{M}}(\mu;\sigma(G(0)))^{d_G} \right] \ \le\ r(X) \quad \forall X\in\Omega_n,$$

where, for any compact $K \subsetneq \mathbb{D}$ and $\mu\in\mathbb{D}$, we define $\mathsf{dist}_{\mathcal{M}}(\mu;K) := \min_{\zeta\in K} \left|(\mu-\zeta)(1-\overline{\zeta}\mu)^{-1}\right|$. For the moment, let us fix $X\in\Omega_n$. For each $\mu\in\sigma(G(X))$, let $\lambda^{(\mu)}$ be an eigenvalue of $G(0)$ such that $\left|(\mu-\lambda^{(\mu)})(1-\overline{\lambda^{(\mu)}}\mu)^{-1}\right| = \mathsf{dist}_{\mathcal{M}}(\mu;\sigma(G(0)))$. Now fix $\mu\in\sigma(G(X))$. The above inequality leads to

$$\left| \frac{\mu-\lambda^{(\mu)}}{1-\overline{\lambda^{(\mu)}}\mu} \right| \ \le\ r(X)^{1/d_G}. \tag{4.1}$$

Applying Lemma 4.1 to the Möbius transformation

$$T(z) \ =\ \frac{|\mu|z-\lambda^{(\mu)}}{1-\overline{\lambda^{(\mu)}}|\mu|z},$$

we deduce that

$$\left| \frac{\zeta-\lambda^{(\mu)}}{1-\overline{\lambda^{(\mu)}}\zeta} \right| \ \ge\ \frac{\big||\mu|-|\lambda^{(\mu)}|\big|}{1-|\mu||\lambda^{(\mu)}|} \quad \forall\zeta : |\zeta| = |\mu|.$$

Applying the above fact to (4.1), we get

$$\begin{aligned} &\frac{|\mu|-|\lambda^{(\mu)}|}{1-|\mu||\lambda^{(\mu)}|} \ \le\ r(X)^{1/d_G} \\ &\Rightarrow \quad |\mu| \ \le \frac{r(X)^{1/d_G}+|\lambda^{(\mu)}|}{1+|\lambda^{(\mu)}|r(X)^{1/d_G}}, \quad \mu\in\sigma(G(X)). \end{aligned} \tag{4.2}$$

Note that the function

$$t \longmapsto \frac{r(X)^{1/d_G}+t}{1+r(X)^{1/d_G}t}, \quad t\ge 0,$$

is an increasing function on $[0,\infty)$. Combining this fact with (4.2), we get

$$|\mu| \ \le\ \frac{r(X)^{1/d_G}+r(G(0))}{1+r(G(0))r(X)^{1/d_G}},$$

which holds $\forall\mu\in\sigma(G(X))$, while the right-hand side is independent of μ. Since this is true for any arbitrary $X\in\Omega_n$, we conclude that

$$r(G(X)) \ \le\ \frac{r(X)^{1/d_G}+r(G(0))}{1+r(G(0))r(X)^{1/d_G}} \quad \forall X\in\Omega_n.$$

In order to prove the sharpness of (1.2), let us fix an $n\ge 2$, and define

$$\mathfrak{S}_n \ :=\ \{A\in\Omega_n : A \text{ has a single eigenvalue of multiplicity } n\}.$$

Pick any $d = 1, \dots, n$, and define

$$M_d(X) \ := \begin{cases} [\mathsf{tr}(X)/n], & \text{if } d = 1, \\ \begin{bmatrix} 0 & & & \mathsf{tr}(X)/n \\ 1 & 0 & & 0 \\ & \ddots & \ddots & \vdots \\ & & 1 & 0 \end{bmatrix}_{d\times d}, & \text{if } d \geq 2, \end{cases}$$

and, for the chosen d, define $\mathsf{G}^{(d)}$ by the following block-diagonal matrix

$$\mathsf{G}^{(d)}(Y) \ := \begin{bmatrix} M_d(X) & \\ & \dfrac{\mathsf{tr}(X)}{n}\mathbb{I}_{n-d} \end{bmatrix} \quad \forall X \in \Omega_n.$$

For our purposes $\mathsf{G}^{A,d} = \mathsf{G}^{(d)}$ for each $A \in \mathfrak{S}_n$; i.e., the equality (1.5) will will hold with the same function for each $A \in \mathfrak{S}_n$. To see this, note that

- $r(\mathsf{G}^{(d)}(X)) = |\mathsf{tr}(X)/n|^{1/d}$; and
- $\mathsf{G}^{(d)}(0)$ is nilpotent of degree d, whence $d_{\mathsf{G}^{(d)}} = d$.

Therefore,

$$\frac{r(A)^{1/d} + r(\mathsf{G}^{(d)}(0))}{1 + r(\mathsf{G}^{(d)}(0))r(A)^{1/d}} \ = \ r(A)^{1/d} \ = \ r(\mathsf{G}^{(d)}(A)) \quad \forall A \in \mathfrak{S}_n,$$

which establishes (1.5) $\square$

References

[1] J. Agler and N.J. Young, *A commutant lifting theorem for a domain in $\mathbb{C}^2$ and spectral interpolation*, J. Funct. Anal. **161** (1999), 452-477.

[2] J. Agler and N.J. Young, *The two-point spectral Nevanlinna-Pick problem*, Integral Equations Operator Theory **37** (2000), 375-385.

[3] J. Agler and N.J. Young, *The two-by-two spectral Nevanlinna-Pick problem*, Trans. Amer. Math. Soc. **356** (2004), 573-585.

[4] J.Agler, N.J. Young, *The hyperbolic geometry of the symmetrized bidisc*, J. Geom. Anal. **14** (2004), 375-403.

[5] H. Bercovici, C. Foias and A. Tannenbaum, *Spectral variants of the Nevanlinna-Pick interpolation problem* in Signal Processing, Scattering and Operator Theory, and Numerical Methods (Amsterdam, 1989), 23-45, Progr. Systems Control Theory **5**, Birkhuser Boston, Boston, MA, 1990.

[6] H. Bercovici, C. Foias, A. Tannenbaum, *A spectral commutant lifting theorem*, Trans. Amer. Math. Soc. **325** (1991), 741-763.

[7] C. Costara, *The 2×2 spectral Nevanlinna-Pick problem*, J. London Math. Soc. (2) **71** (2005), 684-702.

[8] C. Costara, *On the spectral Nevanlinna-Pick problem*, Studia Math. **170** (2005), 23-55.

[9] J. Globevnik, *Schwarz's lemma for the spectral radius*, Rev. Roumaine Math. Pures Appl. **19** (1974), 1009-1012.

[10] M. Jarnicki and P. Pflug, *Invariant Distances and Metrics in Complex Analysis*, de Gruyter Expositions in Mathematics no. 9, Walter de Gruyter & Co., Berlin, 1993.

[11] N. Nikolov, P. Pflug, P.J. Thomas and W. Zwonek, *Estimates of the Carathéodory metric on the symmetrized polydisc*, arXiv preprint `arXiv:math.CV/0608496`.

[12] A. Nokrane and T. Ransford, *Schwarz's lemma for algebroid multifunctions*, Complex Variables Theory Appl. **45** (2001), 183–196.

[13] D.Ogle, *Operator and Function Theory of the Symmetrized Polydisc*, Thesis (1999), `http://www.maths.leeds.ac.uk/ nicholas/`.

[14] T.J. Ransford and M.C. White, *Holomorphic self-maps of the spectral unit ball*, Bull. London Math. Soc. **23** (1991), 256-262.

[15] E. Vesentini, *On the subharmonicity of the spectral radius*, Boll. Un. Mat. Ital. (4) **1** 1968, 427-429.

Department of Mathematics, Indian Institute of Science, Bangalore – 560 012
E-mail address: bharali@math.iisc.ernet.in